



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
-----------------	-------------	----------------------	---------------------	------------------

10/667,465

09/23/2003

Kazuhiro Mochinaga

2003\_1343A

9043

513 7590 09/26/2007  
WENDEROTH, LIND & PONACK, L.L.P.  
2033 K STREET N. W.  
SUITE 800  
WASHINGTON, DC 20006-1021

EXAMINER

SAINT CYR, LEONARD

ART UNIT

PAPER NUMBER

2626

MAIL DATE

DELIVERY MODE

09/26/2007

PAPER

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

# Office Action Summary

Application No.

10/667,465

Applicant(s)

MOCHINAGA ET AL.

Examiner

Leonard Saint-Cyr

Art Unit

2626

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

## Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

## Status

- 1) ☒ Responsive to communication(s) filed on 17 July 2007.
- 2a) ☒ This action is **FINAL**. 2b) ☐ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

## Disposition of Claims

- 4) ☒ Claim(s) 6-13,21-23 and 25-42 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 6-13,21-23 and 25-42 is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

## Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on \_\_\_\_\_ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

## Priority under 35 U.S.C. § 119

- 12) ☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☒ All b) ☐ Some \* c) ☐ None of:
1. ☒ Certified copies of the priority documents have been received.
2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
- \* See the attached detailed Office action for a list of the certified copies not received.

## Attachment(s)

- 1) ☐ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) ☐ Information Disclosure Statement(s) (PTO/SB/08)  
Paper No(s)/Mail Date \_\_\_\_\_
- 4) ☐ Interview Summary (PTO-413)  
Paper No(s)/Mail Date. \_\_\_\_\_
- 5) ☐ Notice of Informal Patent Application
- 6) ☐ Other: \_\_\_\_\_

## DETAILED ACTION

### *Response to Arguments*

1. Applicant's arguments filed 07/17/07 have been fully considered but they are not persuasive.
2. Applicant argues that neither Pitman et al., nor Ellis et al., teach calculating peak value corresponding to values at respective peaks of the band spectra, and obtains, as the prescribed feature quantities, values of difference between peak values of frequency bands (Amendment, pages 19).

The examiner disagrees, Ellis et al., teach utilizing fast Fourier transform process to generate audio frame signatures. Since matches can occur on several consecutive frames, each match (audio and video) has a peak width associated therewith. The number of such consecutively detected matches is referred to as the peak width; examines the run structure in the segment signature and generates an anticipated peak with value therefrom (col.19, lines 25 – 28; col.31, lines 23 – 25; col.45, lines 25 – 30). Generating a peak width value from different frames of an audio signal implies calculating peak value corresponding to values at respective peaks of the band spectra, and obtains, as the prescribed feature quantities, values of difference between peak values of frequency bands, since a peak width is associated with each match in a frame (generated from fast Fourier transform), but not represented the match itself.

3. Applicant argues that neither Pitman et al., nor Ellis et al., teach calculating a cross-correlation value between one of the plurality of signal portions extracted by the signal extracting section and another of the plurality of signal portions, the feature quantity calculating section obtaining a numerical value related to the calculated cross-correlation value as a feature quantity of the audio signal (Amendment, page 19).

The examiner disagrees, Ellis et al., teach producing signatures characterizing respective intervals of a broadcast signal exhibiting correlation between at least some of said respective intervals for use in broadcast segment recognition. The correlator performs the requested matching operation and supplies the match results, along with the relevant information such as corresponding error count (col.5, lines 16 – 19; col.11, lines 8 – 11). Producing signatures characterizing respective intervals of a broadcast signal exhibiting correlation between at least some of said respective intervals, and supplying the match results imply calculating a cross-correlation value between one of the plurality of signal portions extracted by the signal extracting section and another of the plurality of signal portions, since the correlation is calculated between different intervals of a same broadcast signal.

4. Applicant argues that Pitman et al., do not teach derivation of an envelope curve (Amendment, page 20).

The examiner disagrees, Pitman et al., teach rather than looking for a peak in the signal in a frequency channel, another type of curve characteristic such as inflection

point could be used as a trigger event (paragraph 36, lines 1 – 3). Using curve characteristic implies deriving an envelope curve.

***Claim Rejections - 35 USC § 102***

5. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

6. Claims 25 – 30 are rejected under 35 U.S.C. 102(e) as being anticipated by Pitman et al., (US PAP 2002/0143530).

As per claim 25, Pitman et al., teach a feature quantity extracting apparatus comprising: a frequency transforming section for performing a frequency transform on a signal portion corresponding to a prescribed time length, which is contained in an inputted audio signal, to derive frequency spectra from the signal portion (“the audio signal is sampled and a frequency transform is performed on a succession of set of samples”; Abstract, lines 1 –3; paragraph 30, lines 5 - 7);

an envelope curve deriving section for deriving envelope signals which represent envelop curves of the frequency spectra derived by the frequency transforming section (“curve characteristic”; paragraph 32, lines 8 – 11; paragraph 36, lines 1 – 3); and

a feature quantity calculating section for calculating, as feature quantities of the audio signal, numerical values related to respective extremums of the envelope signals derived by the envelope curve deriving section ("local maximum or minimum"; paragraph 35, lines 1 – 3).

As per claim 26, Pitman et al., further disclose that the feature quantity calculating section obtains, as the feature quantities of the audio signal, extremum frequencies each being a frequency corresponding to one of the extremums of the envelope signals derived by the envelope curve deriving section ("an extremum in a semitone frequency channel"; paragraph 35, lines 1 – 3).

As per claim 27, Pitman et al., further disclose that the feature quantity calculating section includes: an extremum frequency calculating section for calculating the extremum frequencies each being a frequency corresponding to one of the extremums of the envelope signals derived by the envelope curve deriving section ("an extremum in a semitone frequency channel"; paragraph 35, lines 1 – 3); and

a space calculating section for calculating spaces between adjacent extremum frequencies as the feature quantities of the audio signal ("evenly spaced frequency band"; paragraph 32).

As per claim 28, Pitman et al., further disclose that the space calculating section obtains, as the feature quantities of the audio signal, numerical values which represent

Art Unit: 2626

a space as a ratio to a prescribed reference value ("equally spaced on a logarithmic scale"; paragraph 32, lines 1 – 4).

As per claims 29, and 30, Pitman et al., further disclose that the space calculating section obtains, as the prescribed reference value, a lowest of the extremum frequencies; a value of difference between a lowest and a second lowest of the extremum frequencies ("evenly spaced frequency band"; paragraph 35, lines 1 – 4; paragraph 32).

### ***Claim Rejections - 35 USC § 103***

7. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

Claims 6 – 13, 21 - 23, and 31 – 42 are rejected under 35 U.S.C. 103(a) as being unpatentable over Pitman et al., (US PAP 2002/0143530) in view Ellis et al., (US Patent 5,504,518).

As per claim 6, Pitman et al., a feature quantity extracting apparatus comprising: a frequency transforming section for performing a frequency transform on a signal portion corresponding to a prescribed time length, which is contained in an inputted audio signal, to derive a frequency spectrum from the signal portion ("the audio signal is

sampled and a frequency transform is performed on a succession of set of samples”;

Abstract, lines 1 –3; paragraph 30, lines 5 - 7);

a band extracting section for extracting a plurality of frequency bands from the frequency spectrum derived by the frequency transforming section and for outputting band spectra which are respective frequency spectra of the extracted frequency bands (“frequency bands”; Abstract, lines 5, and 6; paragraph 32, line 11); and

a feature quantity calculating section for calculating respective prescribed feature quantities of the band spectra, the feature quantity calculating section obtaining the calculated prescribed feature quantities as feature quantities of the audio signal (“extract features from unknown audio content”; paragraph 33; paragraph 54).

However, Pitman et al., do not specifically teach that the feature quantity calculating section calculates peak values corresponding to values at respective peaks of the band spectra, and obtains, as the prescribed feature quantities, values of difference between peak values of frequency bands.

Ellis et al., teach that segment recognition sub-system may detect multiple matches on a given key signature for consecutive frames; examines the run structure in the segment signature and generates an anticipated peak value width therefrom (col.45, lines 25 – 31).

Therefore it would have been obvious to one of ordinary skill in the art at the time the invention was made to generate a peak value among consecutive frames as taught by Ellis et al., in Pitman et al., because that would help better identify the audio content.



As per claim 7, Ellis et al., further disclose the feature quantity calculating section uses binary values to represent the values of difference between peak values of frequency bands, the binary values indicating a sign of a corresponding one of the values of difference ("binary value"; col.15, lines 37 – 40).

As per claim 8, Pitman et al., a feature quantity extracting apparatus comprising: a frequency transforming section for performing a frequency transform on a signal portion corresponding to a prescribed time length, which is contained in an inputted audio signal, to derive a frequency spectrum from the signal portion ("the audio signal is sampled and a frequency transform is performed on a succession of set of samples"; Abstract, lines 1 –3; paragraph 30, lines 5 - 7);

a band extracting section for extracting a plurality of frequency bands from the frequency spectrum derived by the frequency transforming section and for outputting band spectra which are respective frequency spectra of the extracted frequency bands ("frequency bands"; Abstract, lines 5, and 6; paragraph 32, line 11); and

a feature quantity calculating section for calculating respective prescribed feature quantities of the band spectra, the feature quantity calculating section obtaining the calculated prescribed feature quantities as feature quantities of the audio signal ("extract features from unknown audio content"; paragraph 33; paragraph 54).

However, Pitman et al., do not specifically teach that the feature quantity calculating section calculates peak frequencies corresponding to frequencies at

Art Unit: 2626

respective peaks of the band spectra, and obtains, as the prescribed feature quantities, numerical values related to the calculated peak frequencies.

Ellis et al., teach that segment recognition sub-system may detect multiple matches on a given key signature for consecutive frames; examines the run structure in the segment signature and generates an anticipated peak value width therefrom (col.45, lines 25 – 31).

Therefore it would have been obvious to one of ordinary skill in the art at the time the invention was made to generate a peak value among consecutive frames as taught by Ellis et al., in Pitman et al., because that would help better identify the audio content.

As per claim 9, Ellis et al., further disclose that the feature quantity calculating section calculates, as the prescribed feature quantities, values of difference between peak frequencies of frequency bands (“detect multiple matches on a given key signature for consecutive frames, and generate an anticipated peak value”; col.45, lines 25 – 31).

As per claim 10, Ellis et al., further disclose that the feature quantity calculating section represents the prescribed feature quantities using binary values indicating whether a corresponding one of the values of difference between peak frequencies of frequency bands is greater than a prescribed value (“one binary value for positive elements”; col.15, lines 37 – 40).

As per claim 11, Pitman et al., a feature quantity extracting apparatus comprising: a frequency transforming section for performing a frequency transform on a signal portion corresponding to a prescribed time length, which is contained in an inputted audio signal, to derive a frequency spectrum from the signal portion ("the audio signal is sampled and a frequency transform is performed on a succession of set of samples"; Abstract, lines 1 –3; paragraph 30, lines 5 - 7);

a band extracting section for extracting a plurality of frequency bands from the frequency spectrum derived by the frequency transforming section and for outputting band spectra which are respective frequency spectra of the extracted frequency bands ("frequency bands"; Abstract, lines 5, and 6; paragraph 32, line 11); and

a feature quantity calculating section for calculating respective prescribed feature quantities of the band spectra, the feature quantity calculating section obtaining the calculated prescribed feature quantities as feature quantities of the audio signal ("extract features from unknown audio content"; paragraph 33; paragraph 54).

the frequency transforming section extracts from the audio signal the signal portion corresponding to a prescribed time length at prescribed time intervals ("the audio signal is sampled and a frequency transform is performed on a succession of set of samples"; Abstract, lines 1 –3; paragraph 30, lines 5 - 7).

However, Pitman et al., do not specifically teach a peak frequency calculating section for calculating peak frequencies corresponding to frequencies at respective peaks of the band spectra; and a peak frequency time variation calculating section for calculating, as the prescribed feature quantities, numerical values related to respective

time variation quantities of the peak frequencies calculated by the peak frequency calculating section.

Ellis et al., teach that segment recognition sub-system may detect multiple matches on a given key signature for consecutive frames; examines the run structure in the segment signature and generates an anticipated peak value width therefrom (col.45, lines 25 – 31).

Therefore it would have been obvious to one of ordinary skill in the art at the time the invention was made to generate a peak value among consecutive frames as taught by Ellis et al., in Pitman et al., because that would help better identify the audio content.

As per claim 12, Ellis et al., further disclose that the peak frequency time variation calculating section obtains, as the prescribed feature quantities, binary values indicating a sign of a corresponding one of the time variation quantities of the peak frequencies ("binary value"; col.15, lines 37 – 40).

As per claim 13, Ellis et al., further disclose that the peak frequency time variation calculating section obtains, as the prescribed feature quantities, binary values indicating whether a corresponding one of the time variation quantities of the peak frequencies is greater than a prescribed value ("one binary value for positive elements"; col.15, lines 37 – 40).

As per claim 21, Pitman et al., teach a feature quantity extracting apparatus comprising: a signal extracting section for extracting from an extracted audio signal a plurality of signal portions each corresponding to a prescribed time length ("extract features from unknown audio content"; paragraph 54, lines 1 - 10);

However Pitman et al., do not specifically teach calculating a cross-correlation value between one of the plurality of signal portions extracted by the signal extracting section and another of the plurality of signal portions, the feature quantity calculating section obtaining a numerical value related to the calculated cross-correlation value as a feature quantity of the audio signal; the signal extracting section extracts the signal portions at prescribed time intervals, and wherein the feature quantity calculating section includes: a cross-correlation value calculating section for calculating the cross-correlation value at the prescribed time intervals; and a cross-correlation value time variation calculating section for calculating a time variation quantity of the cross-correlation value as the feature quantity of the audio signal.

Ellis et al., teach producing signatures characterizing respective intervals of a broadcast signal exhibiting correlation between at least some of said respective intervals for use in broadcast segment recognition. The correlator performs the requested matching operation and supplies the match results, along with the relevant information such as corresponding error count. (col.5, lines 16 – 19; col.11, lines 8 – 11).

Therefore it would have been obvious to one of ordinary skill in the art at the time the invention was made to determine correlation between some intervals as taught by

Ellis et al., in Pitman et al., because that would help better identify the audio content by supplying matching results to the segment recognition.

As per claim 22, Ellis et al., further disclose that the feature quantity calculating section obtains the cross-correlation value as the feature quantity of the audio signal (“supplies a match report for each audio”; col.11, lines 8 – 13).

As per claim 23, Ellis et al., further disclose that the feature quantity calculating section obtains a binary value as the feature quantity of the audio signal, the binary value indicating a sign of the cross-correlation value (“binary value”; col.15, lines 37 – 40).

As per claims 31 - 33, 35 –37, and 39 –41, Pitman et al., teach a recording medium, and reproduction medium; and a feature quantity storage section which stores at least a set of a feature quantity of an audio signal and control instruction information associated therewith, (paragraph 54; paragraph 25).

However Pitman et al., do not specifically teach receiving television program data containing an audio signal and a video signal, and is capable of recording the television program data to a recording medium, wherein the feature quantity extracting apparatus obtains a feature quantity of the audio signal contained in the television program data, wherein the program recording apparatus further comprises: a recording control section for controlling recording of the television program data to the recording medium; the

audio signal containing music played in a television program to be recorded, the control instruction information instructing the recording control section to perform or stop recording of the television program; a feature quantity comparison section for determining whether the audio signal contained in the television program data matches with the audio signal containing the music played in the television program based on both the feature quantity obtained by the feature quantity extracting apparatus and the feature quantity stored in the feature quantity storage section, and wherein when the feature quantity comparison section determines that the audio signal contained in the television program data matches with the audio signal containing the music played in the television program, the recording control section performs the control of performing or stopping recording of the television program data to the recording medium in accordance with an instruction indicated by control instruction information which is stored in the feature quantity storage section and associated with a feature quantity of the audio signal having been determined as matching with the audio signal containing the music played in the television program.

Ellis et al., teach receiving television broadcast signals over a respective channel and demodulates the received signals to provide baseband video and audio signals. The video and audio signals are thereafter supplied to the segment recognition subsystem wherein frames signatures for each of the video and audio signals are generated which are thereafter compared to store key signatures to determine if a match exists (col.9, lines 55 – 62). The FIR module serve to improve signature stability

by averaging the audio spectral data over a number of television frames, thus to enhance the likelihood of obtaining correct signatures matches (col.21, lines 64 – 67).

Therefore it would have been obvious to one of ordinary skill in the art at the time the invention was made to match audio in a television broadcast as taught by Ellis et al., in Pitman et al., because that would help verify whether video and audio signals are synchronized during television broadcasting.

As per claim 34, 38, and 42, Ellis et al., further disclose that the program reproduction control apparatus further comprises an editing section capable of editing the television program data recorded in the recording medium (updating a broadcast segment recognition database storing signatures”; col.5, lines 2, and 3).

### ***Conclusion***

8. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of



Art Unit: 2626

the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

9. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Leonard Saint-Cyr whose telephone number is (571) 272-4247. The examiner can normally be reached on Mon- Friday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on (571) 272-7602. The fax phone number for the organization where this application or proceeding is assigned is (571)-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

LS  
09/14/07

  
RICHEMOND DORVIL  
SUPERVISORY PATENT EXAMINER